



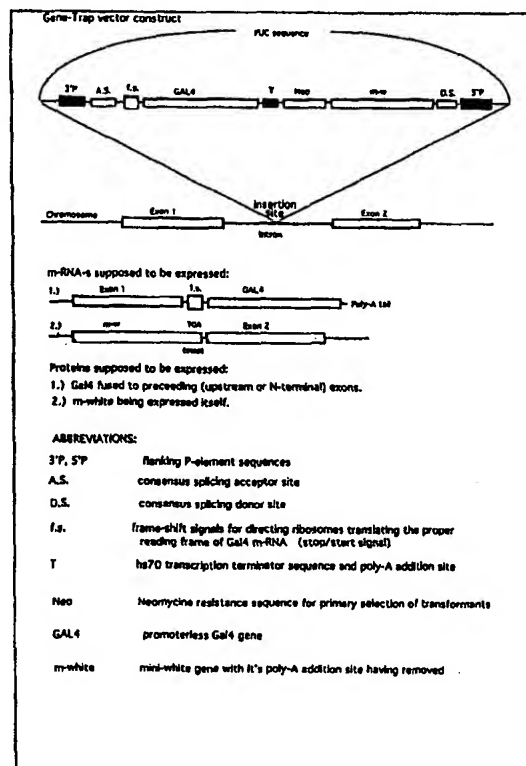
INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12N 15/10, 15/62, 15/90, C12Q 1/68, C07K 14/435		A2	(11) International Publication Number: WO 99/61604
			(43) International Publication Date: 2 December 1999 (02.12.99)
(21) International Application Number: PCT/JP99/02683		(81) Designated States: CA, US, European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).	
(22) International Filing Date: 21 May 1999 (21.05.99)			
(30) Priority Data: 10/141952 22 May 1998 (22.05.98) JP		Published Without international search report and to be republished upon receipt of that report.	
(71) Applicant (for all designated States except US): JAPAN SCIENCE AND TECHNOLOGY CORPORATION [JP/JP]; 1-8, Hon-cho 4-chome, Kawaguchi-shi, Saitama 332-0012 (JP).			
(72) Inventors; and			
(75) Inventors/Applicants (for US only): LUKACSOVICH, Tamas [HU/JP]; 2-30-13, Narusedai, Machida-shi, Tokyo 194-0043 (JP). ASZTALOS, Zoltan [HU/JP]; 3-16-21, Narusedai, Machida-shi, Tokyo 194-0043 (JP). YAMAMOTO, Daisuke [JP/JP]; 4-18-8, Narusedai, Machida-shi, Tokyo 194-0043 (JP). AWANO, Wakae [JP/JP]; Famiyu-102, 3-10-12, Minamidai, Sagami-hara-shi, Kanagawa 228-0814 (JP).			
(74) Agent: NISHIZAWA, Toshio; 6F, Mani-Building, 37-10, Udagawa-cho, Shibuya-ku, Tokyo 150-0042 (JP).			

(54) Title: A VECTOR FOR GENE TRAP, AND A METHOD FOR GENE TRAPPING BY USING THE VECTOR

(57) Abstract

The present application provides a vector for trapping an unknown gene of *Drosophila melanogaster*, which is a recombinant plasmid comprising the following nucleotide sequences in this order: an artificial consensus splicing acceptor site; a synthetic "stop/start" sequence; a reporter gene; a drug resistance gene; a gene responsible for a detectable phenotype of the *Drosophila melanogaster*; and a synthetic splicing donor site. The present application also provides a method for trapping an unknown gene of *Drosophila melanogaster* by using the vector.



FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

Description

A Vector for Gene Trap, and A Method for Gene Trapping
by Using The Vector

5

Technical Field

The present invention relates to a new vector system to facilitate the cloning and functional analysis of new genes
10 of a fly, *Drosophila melanogaster*, and a method for gene trapping with the vector system.

Background Art

There are numerous examples for application of gene
15 trapping methods in wide range of living organisms including maize and mouse (Gossler et al., Science, 244:463-465, 1989).

With respect to tools for gene trapping, the application of different types of enhancer trap P-element vectors (Wilson et al., Genes & Development, 3:1301-1313,
20 1989) for cloning and analyzing trapped genes, as well their use for mosaic analysis with the help of the Gal4/UAS transcription activator system has proven fruitful. However, sometimes the expression pattern of the Gal4 or other reporter gene of the vector construct is affected by
25 enhancers belonging to more than one gene. Similarly, in some cases it is difficult to determine whether the enhancer trap insertion effects the function of one or more of the neighboring genes.

These circumstances altogether with the fact that in
30 some cases the mutant phenotype could be attributed to the

changed expression of a gene with its nearest exon located more than 30 kB apart from the insertion site, can lead in unfortunate cases to an ordeal when it's time to clone and analyze the affected gene.

5 One object of this application is to provide a vector system that includes specifically designed artificial regulatory sequences as well as selection methods for easy screening of positive recombinant lines. More especially, this application intends to provide a vector system of this
10 invention offering much easier and faster cloning opportunities of the affected gene, compared to the widely used enhancer trap P-element vectors. Another object of this application is to provide easier detection method possibilities of the successful trapping events and much
15 higher chance to get more characteristic ("functional") expression patterns of the reporter gene because in the contrary with much of the cases with enhancer trap lines, when using the vector system of this invention, the reporter gene expression is influenced only by a single endogenous
20 transcription unit and effects only the expression of the very same gene.

Disclosure of Invention

The first invention of this application is a vector for
25 trapping an unknown gene of *Drosophila melanogaster*, which is a recombinant plasmid comprising the following nucleotide sequences in this order:

an artificial consensus splicing acceptor site;
a synthetic "stop/start" sequence;
30 a reporter gene;

a drug resistance gene;
a gene responsible for a detectable phenotype of the *Drosophila melanogaster*; and
a synthetic splicing donor site.

5 One embodiment of the first invention is that the recombinant plasmid is derived from pCasper3.

Other embodiments of the first invention are that the reporter gene is the Gal4 gene, Gal4 DNA binding domain-P53 fusion gene or the Gal4-firefly luciferase fusion gene.

10 Further embodiment of this first invention is that the gene responsible for a detectable phenotype of the *Drosophila melanogaster* is mini-white gene.

More further embodiment of the first invention is that the drug resistance gene is neomycin-phosphotransferase gene
15 and its promoter is a heatshock promoter.

The second invention of this application is a method for trapping an unknown gene of *Drosophila melanogaster* by using a vector which is a recombinant plasmid comprising the
20 following nucleotide sequences in this order:

an artificial consensus splicing acceptor site;
a synthetic "stop/start" sequence;
a reporter gene;
a drug resistance gene;
25 a gene responsible for a detectable phenotype of the *Drosophila melanogaster*; and
a synthetic splicing donor site,

which method comprises the steps of:

(a) introducing the vector into the genome of a white minus
30 fly;

- (b) selecting primary transformants resistant to a drug;
- (c) crossing the primary transformants with a transposase source strain to force the vector to jump into other locations;
- 5 (d) selecting secondary transformants by picking up the flies having strong eye color,
- (e) crossing the secondary transformants with UAS (Upstream Activator Sequence)-luciferase harboring strain and measuring the reporter gene expression of the resultant flies; and
- 10 (f) identifying the trapped gene by cloning and sequencing the cDNAs fused to the reporter gene and the gene responsible for a detectable phenotype of the fly.

The third invention of this application is a method for
15 trapping an unknown gene of *Drosophila melanogaster* by using a vector A which is a recombinant plasmid comprising the following nucleotide sequences in this order:

- an artificial consensus splicing acceptor site;
 - a synthetic "stop/start" sequence;
 - 20 Gal4 DNA binding domain-P53 fusion gene as a reporter gene;
 - a drug resistance gene;
 - a gene responsible for a detectable phenotype of the *Drosophila melanogaster*; and
 - a synthetic splicing donor site,
 - 25 and a vector B derived from pCasperhs, which has the heatshock promoter directed Gal4 activator domain-large T antigen fusion gene within polycloning site of the pCasperhs,
- which method comprises the steps of:
- (a) introducing each of the vectors A and B into the
30 genomes of separate white minus flies;

(b) selecting primary transformants for the vector A which are resistant to the drug, and selecting primary transformants for the vector B which have an eye color;

(c) crossing the primary transformants for the vector A with a transposase source strain to force the vector to jump into other locations;

(d) selecting secondary transformants for the vector A by picking up the flies having strong eye color;

(e) crossing the secondary transformants with the primary transformants for the vector B to obtain flies harboring both the vectors A and B;

(f) crossing the flies obtained in the step (e) with an UAS-luciferase harboring fly strain and measuring the reporter gene expression of the resultant flies after a heatshock treatment; and

(g) identifying the trapped gene by cloning and sequencing the cDNAs fused to the reporter gene and the gene responsible for a detectable phenotype of the fly.

Embodiments of the second and third inventions are corresponded to the embodiments of the first invention, and they will be more precisely described in the following description.

Brief Description of Drawings

Figure 1 shows the schematic map of the vector of this invention, pTrap-hsneo.

Figure 2 shows the schematic map of the vector of this invention, pTrap-G4-p53.

Figure 3 shows the schematic map of the vector of this

invention, pCasperhs-G4-LT.

Figure 4 shows the schematic map of the vector of this invention, pTrap-G4-luc.

Figure 5 shows the shematic drawing of a fly genome to
5 which the vector of this invention is inserted for cloning.

Figure 6 shows the results of sequencing RT-PCR products of aop-Gal4 and m-white-aop fusion mRNAs.

Figure 7 presents pictures of characteristic beta-galactosidase staining patterns in different parts of the fly
10 brain resulted from crossing positive gene trap lines with flies harboring a UAS-lacZ construct.

Best Mode for Carrying Out the Invention

A vector construct of the first invention, for example,
15 can be based on the commonly used, P-element transformation vector, pCasper3 (Pirotta, Vectors: A survey of molecular cloning vectors and their uses, eds. Rodriguez, R.L. & Denhardt, D.T., Butterworths, Boston. 437-456, 1998) and the convenient Gal4-UAS expression system (Brand and Perrimon,
20 Development, 118:401-415, 1993).

A promoterless Gal4 gene preceded by an artificial consensus splicing acceptor site and a synthetic "stop/start" sequence to govern the read through translation coming from upstream exon(s) of the trapped gene into the proper reading
25 frame of Gal4 was inserted into the polycloning site of pCasper3.

The removal of the whole 3' UTR (untranslated region) sequence of the mini-white gene and replacement by an artificial splicing donor site resulted in a truncated gene
30 without its own poly-adenylation site.

Without a successful gene trapping event this truncated mini-white gene was not expected to confer any eye color, therefore in this invention a heatshock promoter directed neomycin-phosphotransferase (hs-neo) gene for helping
5 selection of primary transformants by antibiotic feeding has been inserted.

Figure 1 shows the schematic map of the gene trap construct (pTrap-hsneo), and SEQ ID No.1 is the complete nucleotide sequence of the vector pTrap-hsneo.

10 Another gene trap construct, pTrap-G4-p53 (Figure 2) is created by replacing the Gal4 coding sequence of plasmid pTrap-hsneo with a Gal4 DNA binding domain-P53 fusion gene (Clontech, Matchmaker Two Hybrid System, #K1605-1). When this construct coexists in the genome of the same fly with another
15 vector, pCasperhs-G4-LT (Figure 3) containing a heatshock promoter directed Gal4 activator domain-large T antigen (Clontech, Matchmaker Two Hybrid System, #K1605-1) fusion gene, the assembly of a functional Gal4 molecule, through p53-large T antigen interaction, can be regulated by external
20 heatshock.

On this way, the possibility of an intentional temporary control of Gal4 activity became available. In other words the Gal4 expression in a pattern as already determined spatially by the promoter of the trapped gene now can be
25 induced at any desired stage of development by external heatshock.

In order to make the detection of Gal4 expression easier, the Gal4 gene in another construct is replaced with a Gal4-firefly luciferase fusion gene to get pTrap-G4-luc
30 (Figure 4). This artificial gene is coding for a fusion

polypeptide which has preserved both enzymatic activities.

The easy measuring of luciferase activity by luminoassay (Brandes et al., Neuron, 16:687-694, 1996) makes the detection of Gal4 activity comfortable in every single
5 living fly.

Then, one of the best mode of the second or third invention, a method for gene trapping using the vector system, is described in detail.

10 (1) Screening:

The gene trap vector constructs can be introduced into the genome of a white minus fly by microinjection. The selection of primary transformants is possible by using G418, an analog of neomycin, resistance conferred by hs-neo gene.
15 (When performing transformation experiments with these constructs it's turned out that the truncated mini-white gene generally provides a very slight yellow eye color which could be distinguished from w-minus phenotype in most of the cases, therefore G418 selection apparently is not necessary.)

20 After a line with the gene trap construct is being established, the secondary transformants can be generated on the usual way by crossing the original line with a so-called jumpstarter containing the transposase expressing delta 2-3 genetic element.

25 Usually a certain percentage, between 4 and 8, of the secondary transformants have much stronger eye color (deep orange or reddish) than the ancestor fly indicating that the construct was being inserted downstream of a promoter and now the mini-white gene is using the transcriptional "facilities"
30 of that gene (e.g.: poly-adenylation site and transcriptional

terminator) instead of its removed ones. They are the most likely candidates for successful gene trap events. In case of these lines the vector probably has been inserted either into an intron of a gene or upstream from the first intron into the 5' UTR in proper orientation (that is the direction of transcription is same for the "trapped gene" and the mini-white (and Gal4) genes as well). The mini-white gene has its own promoter therefore its expression pattern is supposed to be largely independent from that of the trapped gene.

10 These positive lines are to be checked in the next step for Gal4 expression by crossing them with a "marker" line harboring a UAS-luciferase reporter gene construct. (When using pTrap-G4-luc vector, this step is obviously not necessary.) Usually very strong correlation was found between
15 eye color and Gal4 expression: more than 90% of the lines having strong eye color proved to be expressing Gal4 by means of luciferase assay using luminometer (Brandes et al., Neuron, 16:687-692, 1996).

20 (2) Cloning:

 When the gene trap construct is being inserted into an intron of an endogenous gene, the marker genes of the construct are supposed to be spliced on mRNA level to the exons of the trapped gene by using the artificial splicing
25 acceptor and donor sites. More exactly while the Gal4 mRNA should be joint to the exon(s) located upstream of the insertion site, at the same time the mini-white mRNA is fused to the following exon(s) accomplishing the dual tagging of the trapped gene (Figure 5).

30 This feature can be used for quickly and easily

identifying the trapped gene by means of 3' and 5' RACE (Rapid Amplification of cDNA Ends) experiments. Even cloning and sequencing only a part of the caught mRNA still provides reasonable chance to find homologous mRNAs in the BDGP (Berkeley Drosophila Genome Project) EST (Expressed Sequence Tag) library.

With these approaches, the identification of an already cloned gene can take less than a week compared to the usually more than one year period in average when analyzing a mutant created by some enhancer trap construct.

It's well-known from the literature and the present inventors also have experienced that P-element vectors tend to integrate into or near the 5' UTR of active genes. (The present inventors found that in these cases if the insertion occurred upstream from the first intron, and therefore the artificial splicing acceptor site could not be utilized, the Gal4 gene was expressed by read-through transcription from the nearby promoter.)

The advantage of this tendency can be taken by cloning and sequencing the flanking genomic sequences of the insertion site by inverse or vectorette PCR or by plasmid rescue using suitable restriction digestion to recover the neomycin resistance gene of the construct. Then again the BDGP library can be searched to find any significant matching.

(3) Rescue:

The only reliable way to confirm that any observed mutant phenotype is really the consequence of the P-element insertion is to rescue that particular phenotype. Expectedly the phenotype (some alteration from wild type fly) is caused

by changed expression of gene(s) disturbed by insertion of the P-element. The rescue can be made by expressing the cDNA of the suspected gene most preferable with identical spatial and temporary pattern than that of the gene itself.

5 As it was expected, the vector constructs of the first invention usually cause strong phenotypes. It's not surprising at all because the trapped genes are supposed to be split into two parts on mRNA level resulting in null mutants in majority of the cases. Accordingly mutants
10 obtained by this method frequently show homozygous lethality or sterility. Hypomorphic mutants can be obtained by forcing imprecise excision of the gene trap P-element construct.

As mentioned above, the Gal4 expression is obliged to reflect precisely to that of the trapped gene simply because
15 the Gal4 gene has no its own promoter and they share a common, fused mRNA.

This identical expression provides unique opportunity to rescue the mutant phenotype by crossing this fly with another one harboring the UAS directed, cloned cDNA of the
20 trapped gene.

On this way either the original, homozygous null mutant gene trap fly or any transheterozygous derivative of that with some hypomorphic allele over the null mutant allele can be rescued.

25

(4) Determination of spatial and developmental expression pattern of the trapped gene:

Histochemical determination of the spatially and temporarily controlled expression of any trapped gene is also
30 easy following introduction of a UAS-lacZ construct into the

genome of the same fly and performing either X-gal or antibody staining for beta-galactosidase.

(5) Mosaic analysis:

Possession of a large collection of fly lines with
5 different, characteristic and, in the case of the pTrap-G4-
p53/pCasperhs-G4-TL vector system, inducible Gal4 expression
pattern makes feasible carrying out mosaic analysis of
virtually any gene of interest by directing the expression of
their UAS-constructs on a mutant background with different
10 Gal4 expression patterns.

This approach can answer the question of where and when
that particular gene is required to be expressed to rescue
the mutant phenotype.

Similarly, any gene can be expressed in different
15 ectopic patterns to generate new dominant mutant phenotypes.
This approach might help to conclude the role of that
particular gene and to identify the pathway, in which it's
involved.

20 Example

The following example illustrates a specific embodiment
of the various aspects of the invention. This example is not
intended to limit the invention in any manner.

Figure 6 shows the results of sequencing RT-PCR
25 products of aop-Gal4 and m-white-aop fusion mRNAs.

The template was total RNA prepared from a positive
gene trap line which has the vector pTrap-hsneo being
integrated into the first intron of the well-known aop
(anterior open/pokkuri/yan) developmental gene. The sequences
30 confirm that both splicing occurred precisely at that

particular nucleotides of the artificial regulatory sequences where it was expected.

On Figure 7, there are pictures of characteristic beta-galactosidase staining patterns in different parts of the fly
5 brain resulted from crossing positive gene trap lines with flies harboring a UAS-lacZ construct.

Industrial Applicability

The vector system of this invention offers an
10 exceptional opportunity for easy and fast cloning of the gene responsible for the observed phenotype. Furthermore, by using the UAS-driven coding sequence of any gene of interest, that particular gene can be expressed in identical patterns than those of the trapped genes and these expressions can be
15 regulated temporarily at any desired developmental stage.

Sequence Listing

<110> Japan Science and Technology Corporation
<120> A Vector for Gene Trap, and A Method for Gene Trapping
20 by Using The Vector
<150> Japan, Application No. 10-141952
<151> 22 May 1998
<160> 1
<170> PatentIn Ver. 2.0
25 <210> 1
<211> 11206
<212> DNA
<213> Artificial sequence
<220>
30 <221> 3' P sequence

- <222> (1) .. (237)
<220>
<221> synthetic splicing acceptor site and stop/start
sequence
- 5 <222> (238) .. (274)
<220>
<221> Gal4 gene (coding region and 3' UTR)
<222> (275) .. (3164)
<220>
- 10 <221> hsp70 terminator
<222> (3165) .. (3426)
<220>
<221> synthetic junction sequence
<222> 3427-3457
- 15 <220>
<221> heat shock promoter directed neomycine resistance gene
on complementter strand
<222> (3458) .. (4907)
<220>
- 20 <221> mini-white gene
<222> (4908) .. (8275)
<220>
<221> synthetic splicing donor site
<222> (8276) .. (8299)
- 25 <220>
<221> 5' P sequence
<222> (8300) .. (8446)
<220>
<221> bacterial part of pCasper3 shuttle vector including
- 30 complete pUC8 sequence

<222> (8447) .. (11206)

<220>

<221> synthetic DNA

<222> (238) .. (274)

5 (3427) .. (3457)

(4908) .. (4914)

(8276) .. (8299)

<400> 1

	catgatgaaa taacataagg tggccccgc ggcaagagac atccacttaa cgtatgcttg	60
10	caataagtgc gaggtaaagg aatagtattc tgagtgtcgt attgagtctg agtgagacag	120
	cgatatgatt gttgattaac ccttagcatg tccgtggggg ttgaattaac tcataatatt	180
	aattagacga aattatTTTT aaagTTTTat tTTtaataat ttgcgagtac gcaaagctct	240
	ttctcttaca ggtcgaattg atgtgatgga tccaatgaag ctactgtctt ctatcgaaca	300
	agcatgccat atttgccgac ttaaaaagct caagtgtcc aaagaaaaac cgaagtgcgc	360
15	caagtgtctg aagaacaact gggagtgtcg ctactctccc aaaacccaaa ggtctccgct	420
	gactagggca catctgacag aagtggaaac aaggctagaa agactggaac agctatttct	480
	actgattttt cctcgagaag acctigacat gatTTtgaaa atggattctt tacaggatat	540
	aaaagcattg ttaacaggat tatttgtaca agataatgtg aataaagatg ccgtcacaga	600
	tagattggct tcagtggaga ctgatatgcc tctaacattg agacagcata gaataagtgc	660
20	gacatcatca tcggaagaga gtagtaacaa aggtcaaaga cagttgacig tatcgattga	720
	ctcggcagct catcatgata actccacaat tccgttggat tttatgcccc gggatgctct	780
	tcatggattt gatttggtctg aagaggatga catgtcggat ggcttgcctt tctgaaaac	840
	ggaccccaac aataatgggt tctttggcga cggttctctc ttatgtattc ttogatctat	900
	tggctttaaa ccggaaaatt acacgaactc taacgttaac aggtcccgga ccatgattac	960
25	ggatagatac acgttggctt ctagatccac aacatcccg tttacttcaaa gttatctcaa	1020
	taattttcac ccctactgcc ctatcgtgca ctacccgacg ctaatgatgt tgtataataa	1080
	ccagattgaa atcggtcga aggatcaatg gcaaatcctt tttactgca tattagccat	1140
	tggagcctgg tgtatagagg gggaatctac tgatatagat gtttttact atcaaaatgc	1200
	taaatctcat ttgacgagca aggtcttcga gtcaggttcc ataattttgg tgacagccct	1260
30	acatcttctg tcgcgatata cacagtggag gcagaaaaca aatactagct ataattttca	1320

	cagcttttcc ataagaatgg ccatatcatt gggcttgaat agggacotcc cctcgtcctt	1380
	cagtgatagc agcattctgg aacaaagacg ccgaatttgg tggctctgtct actcttggga	1440
	gatccaattg tccctgcttt atggtcgac catccagctt tctcagaata caatctcctt	1500
	cccttcttct gtgcacgatg tgcagcgtac cacaacaggt cccaccatat atcatggcat	1560
5	cattgaaaca gcaaggctct tacaagtttt cacaaaaatc tatgaactag acaaaacagt	1620
	aactgcagaa aaaagtccta tatgtgcaaa aaaatgcttg atgatttgta atgagattga	1680
	ggaggtttgc agacaggcac caaagttttt acaaatggat atttccacca ccgctctaac	1740
	caatttggtg aaggaacacc ctggctatc ctttacaaga ttcgaaactga agtggaaaca	1800
	gttgctctct atcatttatg tattaagaga ttttttact aattttaccc agaaaaagtc	1860
10	acaactagaa caggatcaaa atgatcatca aagttatgaa gttaaacgat gctccatcat	1920
	gttaagcgat gcagcacaaa gaactgttat gtctgtaagt agctatatgg acaatcataa	1980
	tgtcaoccca tattttgcct ggaattgttc ttattacttg ttcaatgcag tccagtacc	2040
	cataaagact ctactctcaa actcaaaatc gaatgctgag aataacgaga ccgcacaatt	2100
	attacaacaa attaactg ttctgatgct attaaaaaaa ctggccactt ttaaaatcca	2160
15	gacttgtgaa aaatacattc aagtactgga agaggatgt gcgccgtttc tgttatcaca	2220
	gtgtgcaatc ccattaccgc atatcagtta taacaatagt aatggtagcg ccattaaaaa	2280
	tattgtcggc tctgcaacta tcgcccaata cctactctt ccggaggaaa atgtcaacaa	2340
	tatcagtgtt aaatatgttt ctctggctc agtaggcct tcacctgtgc cattgaaatc	2400
	aggagcaagt ttcagtgatc tagtcaagct gttatctaac cgtccaccct ctcgtaactc	2460
20	tccagtgaca ataccaagaa gcacaccttc gcacgctca gtacgcctt ttctagggca	2520
	acagcaacag ctgcaatcat tagtgccact gaccccgctt gctttgtttg gtggcgccaa	2580
	ttttaatcaa agtgggaata ttgctgatag ctcatgttc ttactttca ctaacagtag	2640
	caacggtcog aacctcataa caactcaaac aaattctcaa gcgctttcac aaccaattgc	2700
	ctcctctaac gttcatgata acttcatgaa taatgaaatc acggctagta aaattgatga	2760
25	tggtataaat tcaaaaccac tgtcacctgg ttggacggac caaactgcgt ataacgcgtt	2820
	tggaatcact acagggatgt ttaataccac tacaatggat gatgtatata actatctatt	2880
	cgatgatgaa gatacccac caaacccaaa aaaagagtaa aatgaatcgt agatactgaa	2940
	aaaccccgca agttcacttc aactgtgcat cgtgcaccat ctcaatttct ttcatttata	3000
	catcgttttg ccttctttta tgaactata ctctctaag tttaatctt ggccatgtaa	3060
30	cctctgatct atagaatttt ttaaatgact agaattaatg cccatctttt ttttgacct	3120

	aaattcttca tgaaaatata ttacgagggc ttattcagaa gcttatcgat accgtcgact	3180
	aaagccaaat agaaattatt cagttctggc ttaagttttt aaaagtata ttattttttt	3240
	ggttgtaacc aaccaaaga atgtaataa ctaatacata attatgttag ttttaagtta	3300
	gcaacaaatt gattttagct atattagcta ctgtgttaat aaatagaata tattttattta	3360
5	aagataattc gtttttattg tcagggagtg agtttgctta aaaactcgtt tagatccact	3420
	agaaggaccg cggctcctcg accggatcga aaggagggcg aagaactoca gcatgagatc	3480
	cccgcgctgg aggatcatcc agccggcgtc ccggaacacg attccgaagc ccaacctttc	3540
	atagaaggcg gcggtggaat cgaaatctcg tgatggcagg ttggcgctcg ctgtgtcggt	3600
	catttogaac ccagagtcg cgctcagaag aactcgtcaa gaaggcgata gaaggcgatg	3660
10	cgctcgcaat cgggagcggc gataccgtaa agcacgagga agcggtcagc ccattcgccg	3720
	ccaagctctt cagcaatata acgggtagcc aacgctatgt cctgatagcg gtccgccaca	3780
	cccagccggc cacagtcgat gaatccagaa aagcggccat ttccaccat gatattcggc	3840
	aagcaggcat cgccatgggt cagcagcaga tctcgcctg cgggcagtcg cgccttgagc	3900
	ctggcgaaca gttcggctgg cgcgagcccc tgatgctctt cgtccagatc atcctgatcg	3960
15	acaagaccgg cttccatccg agtacgtgct cgctcgatgc gatgtttcgc ttggtggtcg	4020
	aatgggcagg tagccggatc aagcgtatgc agccgcgca ttgcatcagc catgatggat	4080
	actttctcgg caggagcaag gtgagatgac aggagatcct gccccggcac ttccccaat	4140
	agcagccagt ccttcccg ctcagtgaac acgtcgagca cagctgcgca aggaacgcc	4200
	gtcgtggcca gccacgatag ccgcgtgcc tgcctctgca gttcattcag ggcaccggac	4260
20	aggtcggctt tgacaaaag aaccggcgcc cctgcgctg acagccggaa cacggcggca	4320
	tcagagcagc cgattgtctg ttgtgccag tcatagccga atagcctctc caccgaagcg	4380
	gccggagaac ctgcgtgcaa tccatcttgt tcaatcatgc gaaacgatcc tcatcctgtc	4440
	tcttgatcag atcccctatt cagagttctc ttcttgtatt caataattac ttcttgagc	4500
	atttcagtag ttgcagttga ttacttggt tgctggttac ttttaattga ttactttta	4560
25	cttgcacttt actgcagatt gtttagcttg ttacagtcgc ctgttttatt tgcttagctt	4620
	tcgcttagcg acgtgttcac ttgtctgtt tgaattgaat tgcgctccg tagacgaagc	4680
	gcctctatct atactccggc gctcttttcg cgaacattcg aggcgcgctc tctcgaacca	4740
	acgagagcag tatgccgttt actgtgtgac agagtgcagc agcattagtg cagagaggga	4800
	gagacccaaa aagaaaagag agaataacga ataacggcca gagaaatttc tcgagtttc	4860
30	tttctgcaa acaaatgacc taccacaata accagtttgt ttgggatct agtcccta	4920

	tctagtatgt atgtaagtta ataaaaccct tttttggaga atgtagattt aaaaaaacat	4980
	atTTTTTTTT tatttttttac tgcactggac atcattgaac ttatctgac agttttaaat	5040
	ttacttcgat ccaagggtat ttgaagtacc aggttcttct gattacctct cactcaaaat	5100
	gacattccac tcaaagtcag cgctgtttgc ctcttctct gtccacagaa atatcgccgt	5160
5	ctctttcgcc gctgcgtccg ctatctcttt cgccaccgtt tgtagcgtta cctagcgtca	5220
	atgtccgcct tcagttgcac ttgtcagcg gtttcgtgac gaagctcaa gcggtttacg	5280
	ccatcaatta aacacaaagt gctgtgccaa aactcctctc gcttcttatt ttgtttgtt	5340
	ttttgagtga ttgggtggt gattggtttt ggggtggtaa gcaggggaaa gtgtgaaaaa	5400
	tcccggaat gggccaagag gatcaggagc tattaattcg cggaggcagc aaacacccat	5460
10	ctgccgagca tctgaacaat gtgagtagta catgtgcata catcttaagt tcacttgatc	5520
	tataggaact gcgattgcaa catcaaatig tctgcggcgt gagaactgcg acccacaaaa	5580
	atcccaaacc gcaatgcac aaacaaatag tgacacgaaa cagattattc tggtagctgt	5640
	gctcgctata taagacaatt tttaatgca tatcatgac aagacatcta aaggcattca	5700
	ttttcgacta cattcttttt tacaataaat ataacaacca gatattttaa gctgatocta	5760
15	gatgcacaaa aaataataa aagtataaac ctacttcgta ggatacttcg tttgttcgg	5820
	ggttagatga gcataacgct tgtagtigat atttgagatc ccctatcatt gcagggtgac	5880
	agcggacgct tcgcagagct gcattaacca gggcttcggg caggccaaaa actacggcac	5940
	gctcctgcca cccagtcgc oggaggactc cggttcaggg agcggccaac tagccgagaa	6000
	ccacacctat gcctggcaca atatggacat ctttggggcg gtcaatcagc cgggctccgg	6060
20	atggcggcag ctggtcaacc ggacacgcgg actattctgc aacgagcgac acataccggc	6120
	gccaggaaa catttgctca agaacggtga gtttctattc gcagtcggct gatctgtgtg	6180
	aatcttaat aaagggtcca attaccaatt tgaaactcag ttgcggcgt ggcctatccg	6240
	ggcgaaacttt tggccgtgat gggcagttcc ggtgccgga agacgaccct gctgaatgcc	6300
	cttgccttct gatcgccgca gggcatcaa gtatgccat ccgggatgcg actgctcaat	6360
25	ggccaacctg tggacgcaa ggagatgcag gccagggtcg cctatgtcca gcaggatgac	6420
	ctctttatcg gctocctaac ggccaggga cacctgattt tccaggccat ggtgcggatg	6480
	ccacgacatc tgacctatcg gcagcgagt gcccgcgtgg atcagggtgat ccaggagctt	6540
	tcgctcagca aatgtcagca cacgatcatc ggtgtgccc gcagggtgaa aggtctgtcc	6600
	ggcggagaaa ggaagcgtct ggcattcgcc tccaggcac taaccgatcc gccgttctg	6660
30	atctgcgatg agccacctc cggactggac tcatttaccg cccacagcgt cgtccagggtg	6720

	ctgaagaagc tgtcgcagaa gggcaagacc gtcacctga ccattcatca gccgtcttcc	6780
	gagctgtttg agctctttga caagatcctt ctgatggccg agggcagggg agctttcttg	6840
	ggcactccca gcgaagccgt cgacttcttt tccatgtgag ttcatgtgt ttattaaggg	6900
	tatctagcat tacattacat ctcaactcct atccagcgtg ggtgcccagt gtcctaccaa	6960
5	ctacaatccg gcggactttt acgtacaggt gttggccgtt gtgcccggac gggagatcga	7020
	gtcccgtgat cggatcgcca agatatcgca caattttgct attagcaaag tagcccggga	7080
	tatggagcag ttgttggcca ccaaaaattt ggagaagcca ctggagcagc cggagaatgg	7140
	gtacacctac aaggccacct ggttcatgca gttccgggcg gtctgtggc gatcctggct	7200
	gtcggtgctc aaggaaccac tcctcgtaaa agtgcgactt attcagacaa cggtagtggt	7260
10	ttccagtggg aacaaatgat ataacgtta caattcttgg aaacaaattc gctagatttt	7320
	agttagaatt gcctgattcc acacccttct tagtttttt caatgagatg tatagtttat	7380
	agttttgcag aaaataaata aatttcattt aactcgcgaa catgttgaag atatgaatat	7440
	taatgagatg cgagtaacat ttaattttgc agatggttgc catcttgatt ggccctcatct	7500
	ttttgggcca acaactcaag caagtgggcg tgatgaatat caacggagcc atcttcctct	7560
15	tcctgaccaa catgaccttt caaaacgtct ttgccacgat aaatgtaagt ctgttttaga	7620
	atacatttgc atattaataa ttactaact ttctaataa tcgattcgat ttaggtgttc	7680
	acctcagagc tgccagtttt tatgaggag gcccgagtc gactttatcg ctgtgacaca	7740
	tactttctgg gcaaaacgat tgccgaatta ccgtttttc tcacagtgc actggtcttc	7800
	acggcgattg cctatccgat gatcggactg cggccggag tgctgcactt cttcaactgc	7860
20	ctggcgctgg tcactctggt ggccaatgtg tcaacgtcct tcggatatct aatatcctgc	7920
	gccagctcct cgacctcgat ggcgctgtct gtgggtccgc cggttatcat accattcctg	7980
	ctctttggcg gcttcttctt gaactcgggc tcggtgccag tataacctcaa atggttgcg	8040
	tacctctcat ggttccgtta cgccaacgag ggtctgctga ttaaccaatg ggcggacgtg	8100
	gagccgggcg aaattagctg cacatcgctg aacaccacgt gcccagttc gggcaaggtc	8160
25	atcctggaga cgcttaactt ctccgccgcc gatctgccgc tggactacgt gggctctggcc	8220
	attctcatcg tgagcttccg ggtgctcgca tatctggctc taagacttcg ggcccagcgc	8280
	aaggagtaga aggtaaagtag cggccgcacg taagggttaa tgttttcaa aaaaaattcg	8340
	tcgcacaca acctttcctc tcaacaagca aacgtgcaat gaatttaagt gtatacttcg	8400
	gtaagcttcg gctatcgacg ggaccacctt atgttatctc atcatgggcc agaccacgt	8460
30	agtccagcgg cagatcggcg gcggagaagt taagcgtctc caggatgacc ttgccgaac	8520

	tggggcacgt ggtgttcgac gatgtgcagc taatttcgcc cggctccacg tccgccatt	8580
	ggttaatcag cagaccctcg ttggcgtaac ggaaccatga gaggtacgac aaccatttga	8640
	ggtatactgg caccgagccc gagttcaaga agaaggcgtt tttccatagg ctccgcccc	8700
	ctgacgagca tcacaaaaat cgacgctcaa gtcagagggtg gcgaaacccc acaggactat	8760
5	aaagatacca ggcgtttccc cctggaagct cctcgtgcg ctctcctgtt ccgaccctgc	8820
	cgcttacogg atacctgtcc gcctttctcc cticgggaag cgtggcgctt tctcaatgct	8880
	cacgctgtag gtatctcagt tcggtgtagg tcgttcgctc caagctgggc tgtgtgcacg	8940
	aacccccgt tcagcccgac cgctgcgcct tatccggtaa ctatcgtctt gagtccaacc	9000
	cggtaagaca cgacttatcg ccactggcag cagccactgg taacaggatt agcagagcga	9060
10	ggtatgtagg cgggtgtaca gagttcttga agtgggtggc taactacggc taaactagaa	9120
	ggacagtatt tggatatctgc gctctgtga agccagttac cticggaaaa agagtggta	9180
	gctcttgatc cggcaaacaa accaccgctg gtagcgggtg ttttttgtt tgcaagcagc	9240
	agattacgcg cagaaaaaaaa ggatctcaag aagatccttt gatcttttct acggggtctg	9300
	acgctcagtg gaacgaaaac tcacgttaag ggatttttgt catgagatta tcaaaaagga	9360
15	tcttcaccta gatcctttta aattaaaaat gaagttttaa atcaatctaa agtatatatg	9420
	agtaaaacttg gtctgacagt taccaatgct taatcagtga ggcacctatc tcagcgatct	9480
	gtctatttcg ttcattccata gttgcctgac tcccgcgtg gtagataact acgatacggg	9540
	agggcttacc atctggcccc agtgtgcaa tgataccgag agaccacgc tcaccgctc	9600
	cagatttatc agcaataaac cagccagccg gaagggccga gcgcagaagt ggtcctgcaa	9660
20	ctttatccgc ctccatccag tctattaatt gttgccggga agctagagta agtagttgc	9720
	cagttaatag ttgcgcaac gttgttgcca ttgctacagg catcgtggtg tcacgctcgt	9780
	cgtttggtat ggcttcattc agctccggtt cccaacgac aaggcgagtt acatgatccc	9840
	ccatgttgtg caaaaaagcg gttagctcct tcggtcctcc gatcgttgtc agaagtaagt	9900
	tggccgcagt gttatcactc atggttatgg cagcaactgca taattctctt actgtcatgc	9960
25	catccgtaag atgcttttct gtgactgggt agtactcaac caagtcattc tgagaatagt	10020
	glatcggcg accgagttgc tcttgcccg cgtcaacacg ggataatacc gcgccacata	10080
	gcagaacttt aaaagtgtc atcattggaa aacgttcttc ggggcgaaaa ctctcaagga	10140
	tcttaccgct gttgagatcc agttcagatg aaccactcg tgcacccaac tgatcttcag	10200
	catcttttac ttccaccagc gtttctgggt gagcaaaaac aggaaggcaa aatgccgcaa	10260
30	aaaagggaat aaggcgaca cggaaatgtt gaatactcat actcttctt tttcaatatt	10320

attgaagcat ttatcagggt tattgtctca tgagcggata catatttgaa tgtatttaga 10380
aaaataaaca aataggggtt cgcgcacat ttccccgaaa agtgccacct gacgtctaag 10440
aaaccattat tatcatgaca ttaacctata aaaataggcg tatcacgagg ccctttcgtc 10500
tcgcgcgttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca 10560
5 cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 10620
tiggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 10680
accatatgcg gtgtgaaata ccgcaccgaa tcgcgcggaa ctaacgacag tcgctccaag 10740
gtcgtcgaac aaaagtgaa tgtgttcggg agagcgggtg ggagacagcg aaagagcaac 10800
tacgaaacgt ggtgtggtgg aggtgaatta tgaagaggcg gcgcgatttg aaaagtatgt 10860
10 atataaaaaa tatatcccg tgttttatgt agcgataaac gagtttttga tgtaaggtat 10920
gcaggtgtgt aagtcttttg gttagaagac aaatccaaag tctacttggt gggatgttcg 10980
aaggggaaat acttgtattc tataggtcat atcttgtttt tattggcaca aatataatta 11040
cattagcttt ttgagggggc aataaacagt aaacacgatg gtaataatgg taataaaaaa 11100
aacaagcagt tatttcggat atatgtcggc tactccttgc gtcgggcccc aagtcttaga 11160
15 gccagatatg cgagcacccg gaagctcacg atgagaatgg ccagac 11206

CLAIMS

1. A vector for trapping an unknown gene of *Drosophila melanogaster*, which is a recombinant plasmid comprising the following nucleotide sequences in this order:
 - an artificial consensus splicing acceptor site;
 - a synthetic "stop/start" sequence;
 - a reporter gene;
 - a drug resistance gene;
 - a gene responsible for a detectable phenotype of the *Drosophila melanogaster*; and
 - a synthetic splicing donor site.
2. The vector of claim 1, wherein the recombinant plasmid is derived from pCasper3.
3. The vector of claim 1 or 2, wherein the reporter gene is the Gal4 gene.
4. The vector of claim 3, which has the nucleotide sequence of SEQ ID No. 1.
5. The vector of claim 1 or 2, wherein the reporter gene is Gal4 DNA binding domain-P53 fusion gene.
6. The vector of claim 1 or 2, wherein the reporter gene is the Gal4-firefly luciferase fusion gene.
7. The vector of any one of claims 1-6, wherein the gene responsible for a detectable phenotype of the *Drosophila melanogaster* is mini-white gene.

8. The vector of any one of claims 1-7, wherein the drug resistance gene is neomycin-phosphotransferase gene and its promoter is a heatshock promoter.

5

9 A vector derived from pCasperhs, which has the heatshock promoter directed Gal4 activator domain-large T antigen fusion gene within polycloning site of the pCasperhs.

10 10. A method for trapping an unknown gene of *Drosophila melanogaster* by using a vector which is a recombinant plasmid comprising the following nucleotide sequences in this order:

an artificial consensus splicing acceptor site;

a synthetic "stop/start" sequence;

15 a reporter gene;

a drug resistance gene;

a gene responsible for a detectable phenotype of the *Drosophila melanogaster*; and

a synthetic splicing donor site,

20 which method comprises the steps of:

(a) introducing the vector into the genome of a white minus fly;

(b) selecting primary transformants resistant to a drug;

(c) crossing the primary transformants with a transposase
25 source strain to force the vector to jump into other locations;

(d) selecting secondary transformants by picking up the flies having strong eye color,

(e) crossing the secondary transformants with UAS (Upstream
30 Activator Sequence)-luciferase harboring strain and measuring

the reporter gene expression of the resultant flies; and

(f) identifying the trapped gene by cloning and sequencing the cDNAs fused to the reporter gene and the gene responsible for a detectable phenotype of the fly.

5

11. The method according to claim 10, wherein the recombinant plasmid is derived from pCasper3.

12. The method according to claim 10 or 11, wherein the
10 reporter gene in the vector is the Gal4 gene, and in the step (e) the Gal4 expression is measured.

13. The method according to claim 10 or 11, wherein the
15 reporter gene of the vector is the Gal4-firefly luciferase fusion gene, and in the step (e) expression of said fusion gene is measured without crossing the secondary transformants with UAS-luciferase harboring strain.

14. The method according to any one of claims 10 to 14,
20 wherein the gene responsible for a detectable phenotype of the *Drosophila melanogaster* is mini-white gene, and in the step (f) the cDNAs fused to the reporter gene and the mini-white gene are cloned and sequenced.

15. The method according to any one of claims 10 to 15,
25 wherein the drug resistance gene is neomycin-phosphotransferase gene and its promoter is a heatshock promoter, and in the step (b) the transformants resistant to G418 is selected.

30 16. A method for trapping an unknown gene of *Drosophila*

melanogaster by using a vector A which is a recombinant plasmid comprising the following nucleotide sequences in this order:

- an artificial consensus splicing acceptor site;
 - 5 a synthetic "stop/start" sequence;
 - Gal4 DNA binding domain-P53 fusion gene as a reporter gene;
 - a drug resistance gene;
 - a gene responsible for a detectable phenotype of the *Drosophila melanogaster*; and
 - 10 a synthetic splicing donor site,
- and a vector B derived from pCasperhs, which has the heatshock promoter directed Gal4 activator domain-large T antigen fusion gene within polycloning site of the pCasperhs, which method comprises the steps of:
- 15 (a) introducing each of the vectors A and B into the genomes of separate white minus flies;
 - (b) selecting primary transformants for the vector A which are resistant to a drug, and selecting primary transformants for the vector B which have an eye color;
 - 20 (c) crossing the primary transformants for the vector A with a transposase source strain to force the vector to jump into other locations;
 - (d) selecting secondary transformants for the vector A by picking up the flies having strong eye color;
 - 25 (e) crossing the secondary transformants with the primary transformants for the vector B to obtain flies harboring both the vectors A and B;
 - (f) crossing the flies obtained in the step (e) with an UAS-luciferase harboring fly strain and measuring the
 - 30 reporter gene expression of the resultant flies after a

heatshock treatment; and

(g) identifying the trapped gene by cloning and sequencing the cDNAs fused to the reporter gene and the gene responsible for a detectable phenotype of the fly.

5

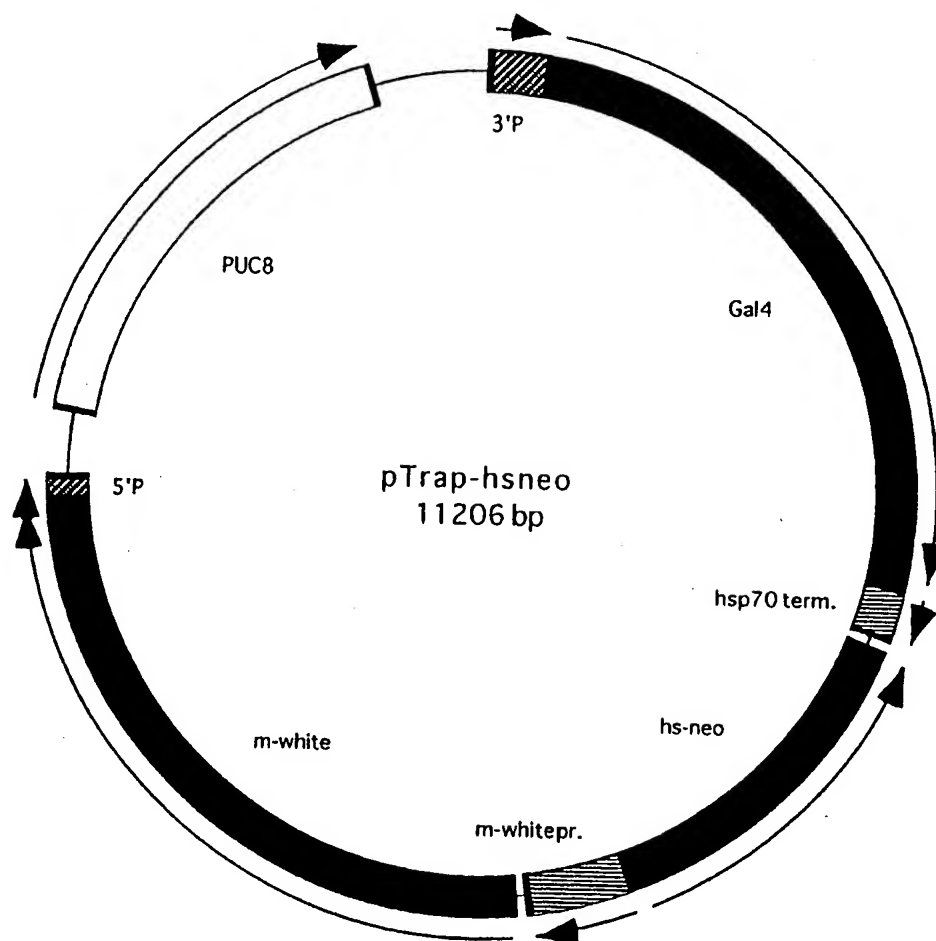
17. The method according to claim 16, wherein the vector A is derived from pCasper3.

10 18. The method according to claim 16 or 17, wherein the gene responsible for a detectable phenotype of the *Drosophila melanogaster* is mini-white gene, and in the step (g) the cDNAs fused to the reporter gene and the mini-white gene are cloned and sequenced.

15 19. The method according to any one of claims 16 to 18, wherein the drug resistance gene is neomycin-phosphotransferase gene and its promoter is a heatshock promoter, and in the step (b) the transformant resistant to G418 is selected.

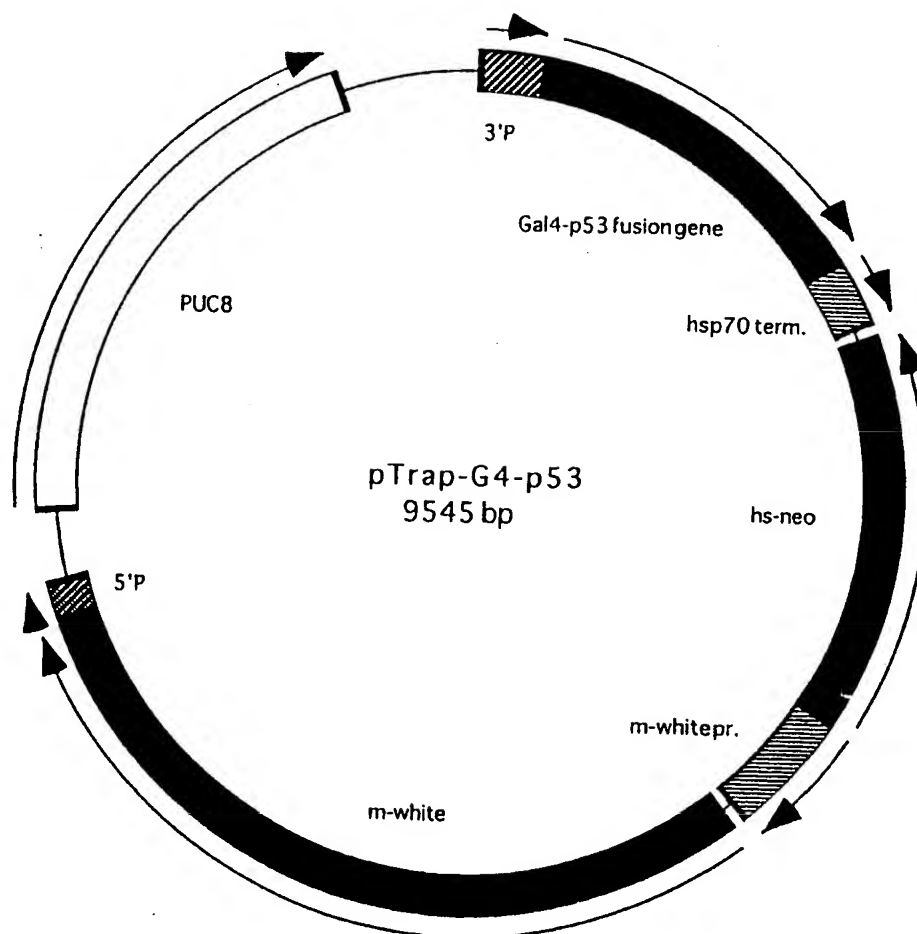
1/7

Fig. 1



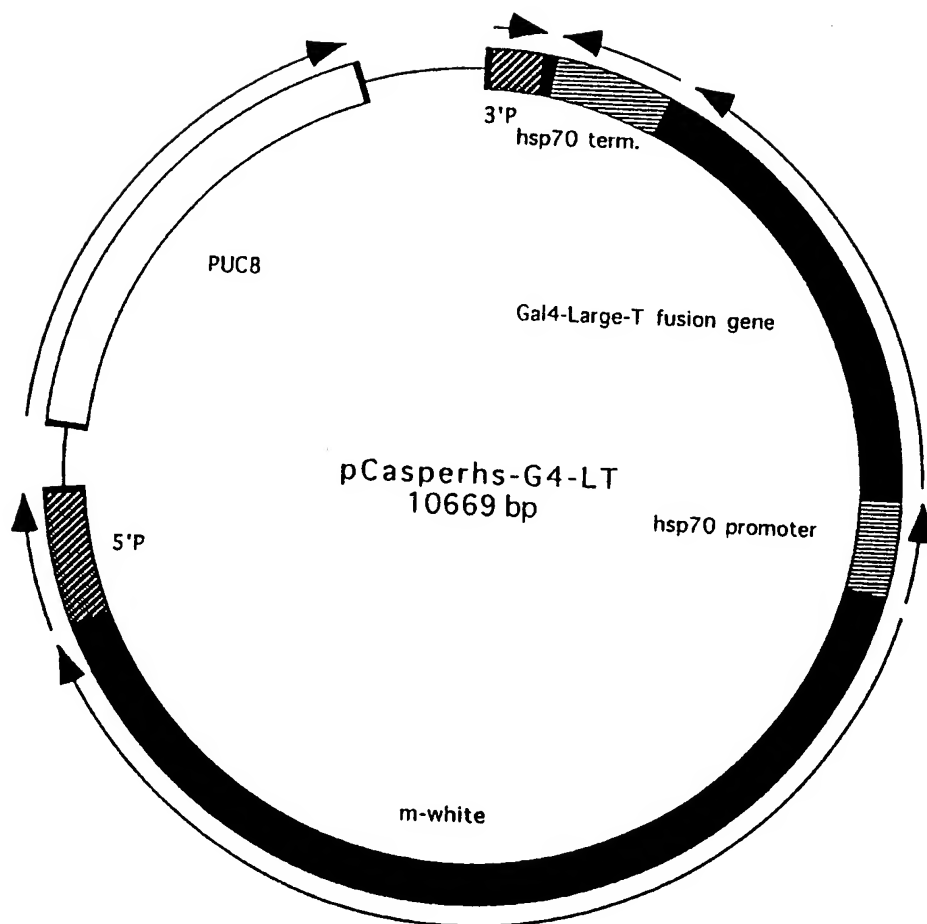
2/7

Fig. 2



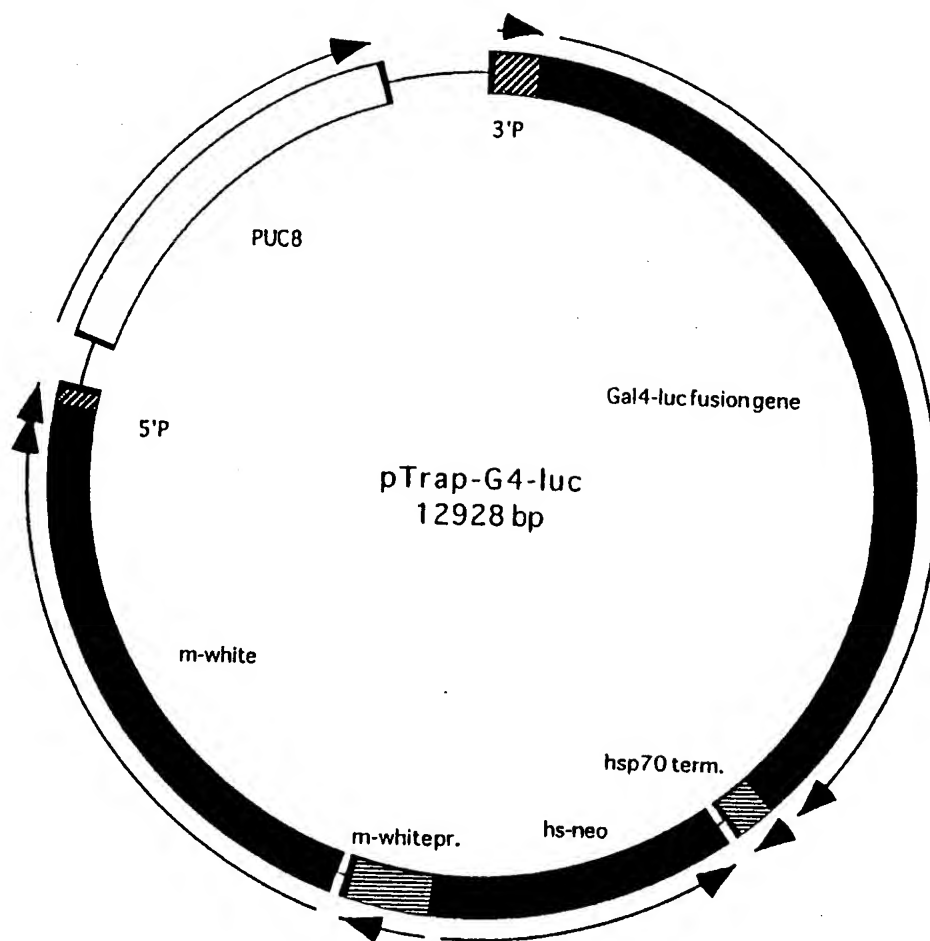
3/7

Fig. 3



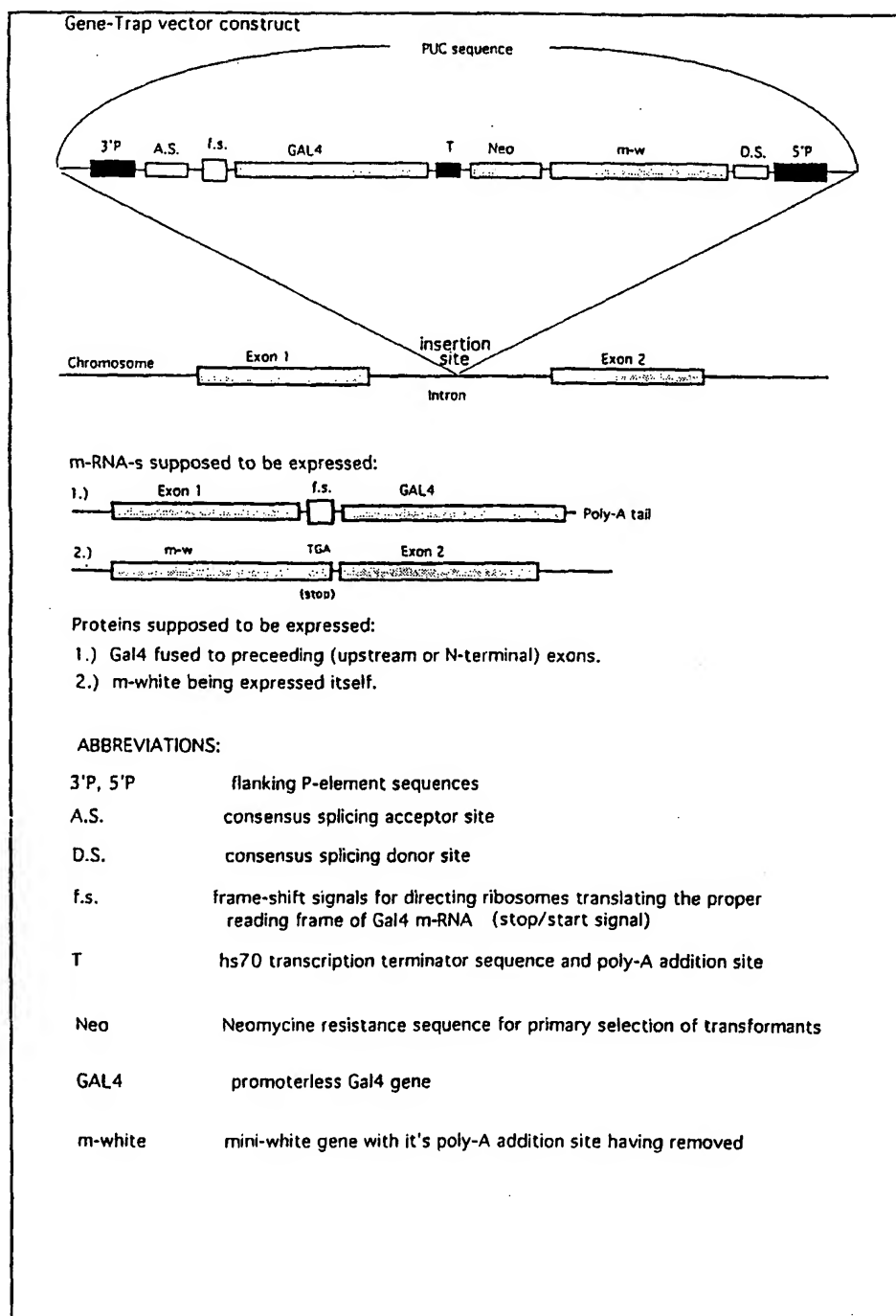
4/7

Fig. 4



5/7

Fig. 5



6/7

Fig. 6

Precise splicing of Gal4 and mini-white genes from Gene Trap vector into anterior open gene

5'P end of vector; splice acceptor site/stop-start seq.;Gal4 gene;mini-white gene/splice donor site; 3'P end of *GT* vector

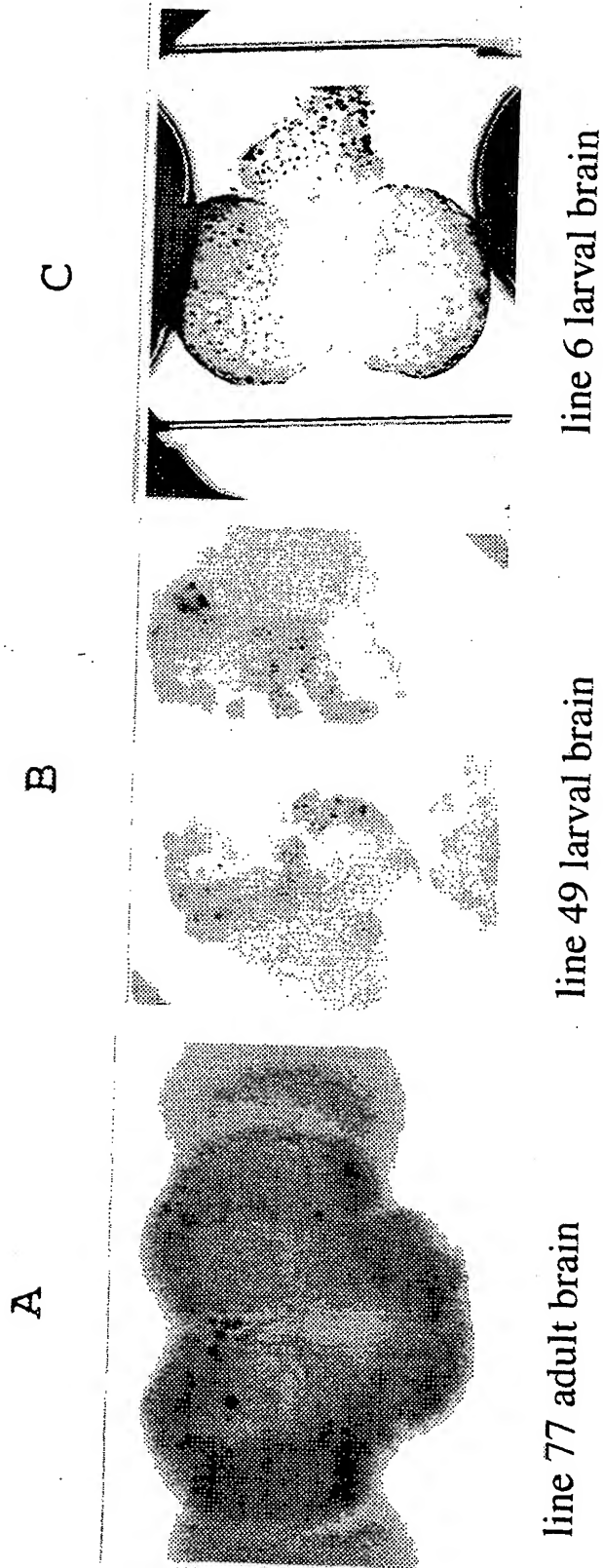
anterior open cDNA exon 1 exon 2
cagtagtaaccacagtgcacagtggagcaaacag/atatacccgcttanaagcaattccagtgygtttccac

anterior open exon 1 - Gal4 fusion cDNA

mini-white - anterior open exon 2 fusion cDNA

Gal4 expression patterns revealed by UAS-lacZ reporter construct.

Fig. 7



PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau



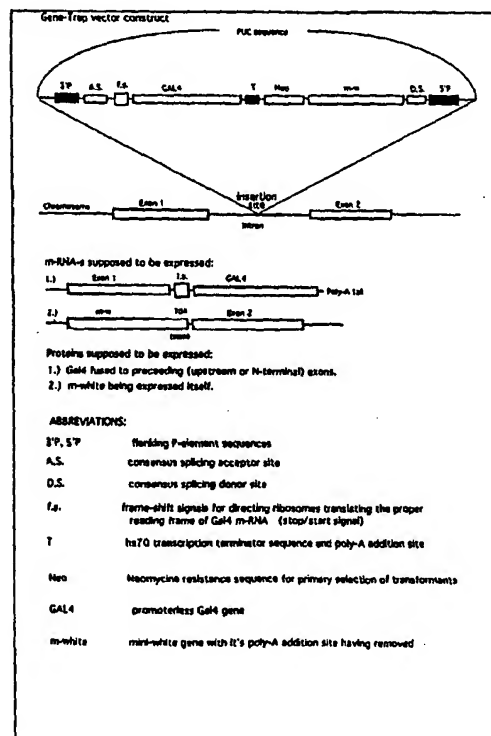
INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ : C12N 15/10, 15/62, 15/85, 15/90, C12Q 1/68, C07K 14/435	A3	(11) International Publication Number: WO 99/61604 (43) International Publication Date: 2 December 1999 (02.12.99)
(21) International Application Number: PCT/JP99/02683 (22) International Filing Date: 21 May 1999 (21.05.99) (30) Priority Data: 10/141952 22 May 1998 (22.05.98) JP (71) Applicant (for all designated States except US): JAPAN SCIENCE AND TECHNOLOGY CORPORATION [JP/JP]; 1-8, Hon-cho 4-chome, Kawaguchi-shi, Saitama 332-0012 (JP). (72) Inventors; and (75) Inventors/Applicants (for US only): LUKACSOVICH, Tamas [HU/JP]; 2-30-13, Narusedai, Machida-shi, Tokyo 194-0043 (JP). ASZTALOS, Zoltan [HU/JP]; 3-16-21, Narusedai, Machida-shi, Tokyo 194-0043 (JP). YAMAMOTO, Daisuke [JP/JP]; 4-18-8, Narusedai, Machida-shi, Tokyo 194-0043 (JP). AWANO, Wakae [JP/JP]; Famiyu-102, 3-10-12, Minamidai, Sagami- hara-shi, Kanagawa 228-0814 (JP). (74) Agent: NISHIZAWA, Toshio; 6F, Mani-Building, 37-10, Udagawa-cho, Shibuya-ku, Tokyo 150-0042 (JP).	(81) Designated States: CA, US, European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE). Published With international search report. (88) Date of publication of the international search report: 2 March 2000 (02.03.00)	

(54) Title: A VECTOR FOR GENE TRAP, AND A METHOD FOR GENE TRAPPING BY USING THE VECTOR

(57) Abstract

The present application provides a vector for trapping an unknown gene of *Drosophila melanogaster*, which is a recombinant plasmid comprising the following nucleotide sequences in this order: an artificial consensus splicing acceptor site; a synthetic "stop/start" sequence; a reporter gene; a drug resistance gene; a gene responsible for a detectable phenotype of the *Drosophila melanogaster*; and a synthetic splicing donor site. The present application also provides a method for trapping an unknown gene of *Drosophila melanogaster* by using the vector.



FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CC	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

INTERNATIONAL SEARCH REPORT

International Application No

PCT/JP 99/02683

A. CLASSIFICATION OF SUBJECT MATTER

IPC 6 C12N15/10 C12N15/62 C12N15/85 C12N15/90 C12Q1/68
C07K14/435

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 6 C12N C12Q C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 98 20031 A (JARVIK JONATHAN W) 14 May 1998 (1998-05-14) the whole document ---	1-19
A	A.H. BRAND AND N. PERRIMON: "Targeted gene expression as a means of altering cell fates and generating dominant phenotypes" DEVELOPMENT, vol. 118, 1993, pages 401-415, XP000857179 THE COMPANY OF BIOLOGISTS, LIMITED, CAMBRIDGE, GREAT BRITAIN cited in the application the whole document --- -/-	1-19

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

Special categories of cited documents:

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

- "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- "&" document member of the same patent family

Date of the actual completion of the international search

3 December 1999

Date of mailing of the international search report

22/12/1999

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Authorized officer

Hornig, H

INTERNATIONAL SEARCH REPORT

International Application No
PCT/JP 99/02683

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	<p>W. WURST ET AL.: "A large-scale gene trap screen for insertional mutations in developmental mutations in developmentally regulated genes in mice"</p> <p>GENETICS, vol. 139, no. 2, February 1995 (1995-02), pages 889-899, XP000857167 GENETIC SOCIETY OF AMERICA, BALTIMORE, MD, US the whole document</p> <p style="text-align: center;">---</p>	1-19
A	<p>C. WILSON ET AL.: "P-element-mediated enhancer detection: an efficient method for isolating and characterizing developmentally regulated genes in Drosophila"</p> <p>GENES & DEVELOPMENT, vol. 3, no. 9, September 1989 (1989-09), pages 1301-1313, XP000857178 CSH LABORATORY PRESS, NEW YORK, US cited in the application the whole document</p> <p style="text-align: center;">---</p>	1-19
A	<p>P. BARTHMAIER AND E. FYRBERG: "Monitoring development and pathology of Drosophila indirect flight muscles using green fluorescent protein"</p> <p>DEVELOPMENTAL BIOLOGY, vol. 169, no. 2, June 1995 (1995-06), pages 770-774, XP002124662 ACADEMIC PRESS, INC., US the whole document</p> <p style="text-align: center;">---</p>	1-19
A	<p>C.S. THUMMEL ET AL.: "Vectors for Drosophila P-element-mediated transformation and tissue culture transfection"</p> <p>GENE, vol. 74, 1988, pages 445-456, XP002124663 ELSEVIER SCIENCE PUBLISHERS, B.V., AMSTERDAM, NL; the whole document</p> <p style="text-align: center;">---</p>	
A	<p>V. PIRROTTA: "Vectors for P-mediated transformation in Drosophila"</p> <p>BIOTECHNOLOGY, VECTORS A SURVEY OF MOLECULAR CLONING VECTORS AND THEIR USES; R. L. RODRIGUEZ AND D.T. DENHARDT, vol. 1, 1988, pages 437-456, XP000857168 Butterworths, Boston, US cited in the application the whole document</p> <p style="text-align: center;">---</p>	
	-/--	

INTERNATIONAL SEARCH REPORT

International Application No
PCT/JP 99/02683

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	<p>C.S. THUMMEL AND V. PIRROTTA: "New pCasPer P element vectors" EMBL SEQUENCE DATABASE, 23 July 1996 (1996-07-23), XP002124664 Cambridge, UK Accession no. EMSYN.PEU59055; U59055; & DROS. INFO. SERVICE, vol. 71, 1992, page 150 ---</p>	
A	<p>A. GOSSLER ET AL.: "Mouse embryonic stem cells and reporter constructs to detect developmentally regulated genes" SCIENCE, vol. 244, 1989, pages 463-465, XP002124665 AAAS, WASHINGTON, DC, US cited in the application the whole document -----</p>	

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/JP 99/02683

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9820031 A	14-05-1998	AU 5168598 A	29-05-1998